To appear in: GWAI-84, J. Laubsch (ed.), Springer, Heidelberg.

NASA-CR-193204

COGNITION AND REPRESENTATION

An overview of knowledge representation issues in cognitive science*

Christian Freksa
Institut für Informatik
TU München
Arcisstr. 21
8000 München 2

Ulrich Furbach
Fachbereich Informatik
HSBw München
Werner-Heisenberg-Weg 39
8014 Neubiberg

Gerhard Dirlich
Abteilung Biostatistik
MPI für Psychiatrie
Kraepelinstr. 2
8000 München 40

ABSTRACT

Central topics and scope of cognitive science are outlined. The role of representations in cognitive systems and metaphors for cognition are discussed. A representation-theoretical approach to knowledge representation is proposed. The notion of a representation system is applied as a framework for specifying, comparing, and transforming representations. We show how this framework can be used to study properties of representations and to make some of the controversies in knowledge representation appear less controversial. Advantages and difficulties of a representation-theoretical approach to cognition are presented.

(NASA-CR-193204) COGNITION AND REPRESENTATION: AN OVERVIEW OF KNOWLEDGE REPRESENTATION ISSUES IN COGNITIVE SCIENCE (Technische Univ.) 26 p

N93-72367

Unclas

Z9/63 0161799

1 COGNITIVE SCIENCE

Cognitive science is concerned with the 'mechanism' of the mind. The function of the mind has been of central interest in psychology since its beginning as a scientific discipline. Workers in artificial intelligence have attempted to construct working models of thought. In the 1970s researchers in philosophy, psychology, artificial intelligence, linguistics, computer science, anthropology, and the neurosciences realized that they might be investigating related problems and that these individual disciplines might profit from interrelating their work. In 1977 the Cognitive Science Society was founded, the journal *Cognitive Science* was established and in 1978, the Sloan Foundation in the United States took the first steps to establish research centers for cognitive science. Since then, the field has attracted growing attention and something like a 'cognitive paradigm' has entered a number of areas of research, development, and application.

* Research supported, in part, by NASA cooperative agreement NCC2-275 and by ESPRIT LOKI pilot project 107.

What were the driving forces that gave cognitive science a chance to emerge between the clearly marked territories of psychology and artificial intelligence? Morton Hunt (1982) notes that in the 1960s and -70s there were tendencies to believe in an inferiority of the role of thinking as compared to the biological basis of the organism in ethology, of the role of learned stimulus-response patterns in psychology, and of the role of emotion. Hunt is convinced that these tendencies resulted in a dissatisfaction which supported the evolution of cognitive science.

Another factor that has played an important role in the development of cognitive science is a paradigm shift in cognitive psychology that took place in recent years. Interest is no longer exclusively focused on high-level mental functions like playing chess, solving math problems, etc. but has also turned to low-level cognitive abilities like creating percepts from visual input or recognizing spoken language. This paradigm shift has helped cognitive science to emerge outside artificial intelligence, whose domain had been higher-level thinking processes for quite some time.

1.1 What is cognition?

From our own experience we all are familiar with the wide spectrum of functions of our mind. These functions range from perceptual processes close to our sense organs to conscious and subconscious decision processes, among them the great variety of high-level intelligent processes, to mental processes that control action. Some examples that have heen investigated in cognitive science are:

- how do we read and understand words like the word "word"?
- how do we know whether the door to our apartment opens left or right?
- how do we understand jokes?
- how do we know how an electric circuit works?
- how do we know how to play tennis?

Cognitive science is concerned with a better understanding of cognitive processes in natural, artificial, and hybride systems. A common research objective for the various subdisciplines of cognitive science is to discover the representational and computational capacities of the mind and their structural and functional representation in the brain. Besides presenting a scientific challenge in its own right, a better understanding of thinking, teaching, and learning processes could aid in the development of systems which complement human cognitive performance in a constructive way.

1.2 The relation of cognitive science to other scientific areas

The cognitive science perspective is a result of the interaction of various disciplines in connection with the rapid developments in computer technology.

- In psychology, the easy-to-understand structure of computers exhibiting remarkably complex behavior has stimulated the introduction of the information processing metaphor, which initiated new theoretical and empirical research in psychology (c.f. Newell and Simon 1972). Examples are given in sections 1.3 and 2.2.

- Syntactic descriptions of programming languages and attempts to model natural language by computer brought together linguistics and computer science (see also the report on the panel discussion on AI and linguistics in this volume).

- Observations of culture-dependent perception, language, and behavior in anthropology are being investigated with methods developed in artificial intelligence (Kay 1981).

- Considerations about the nature of knowledge and intelligence developed in philosophy could be probed within a new testbed (Dreyfus 1979).

- Laws of thought postulated in logic could be implemented on computers and led to the development of new programming languages (Kowalski 1974, Weyhrauch 1980).

- Great difficulties in building scene understanding systems attracted computer scientists' interest in the anatomy and physiology of biological visual systems (Marr 1982). Great difficulties in combining the vast amount of results from the neurosciences attracted neuroscientists' interest in simulations that became possible with methods developed in AI.

Researchers in these various areas spoke different languages and used different paradigms, but an uncertain feeling developed that they all really were tackling the same problem. They all wanted to understand *what thinking is all about*. In this respect, cognitive science can be viewed as an umbrella discipline for the individual subfields.

1.3 Example of research in cognitive science

Examples of research activities in cognitive science are the study of human perception and memory, the scientific foundation of the design of human/machine systems, the study of language use and understanding, the modelling of cognitive processes using methods from artificial intelligence, and the investigation of the nature of 'intelligence'.

A specific example showing the possibility for interaction between psychology and AI is Geoffrey Hinton's (1980) cube problem: imagine a cube suspended (by a string) on one of its corners (say corner 'A') such that the most distant corner points down vertically (compare Fig.1):

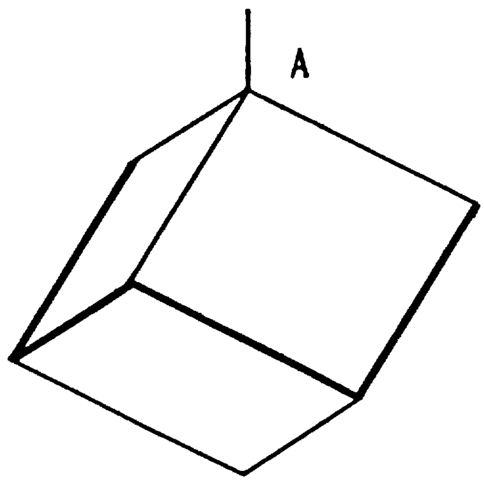


Fig.1: Hinton's cube viewed from the side

you will see a chain of edges zig-zaging around the cube (depicted in the figure by a dark line). Now imagine, you want to turn the cube around the vertical axis in such a way that its topography is preserved, i.e., that each corner of the cube coincides with a corner and each edge with an edge in the original position. By how many degrees do you have to turn the cube to obtain this coincidence for the first time?

For most people it is rather difficult to find the correct solution to this problem. If, however, the problem is presented differently, its solution becomes much easier. Imagine looking along the vertical rotation axis of the cube at corner 'A'. You will see three sides of the cube point-symmetrically arranged around corner 'A' and will recognize instantly that a turn by 120° will preserve the cube's topography.

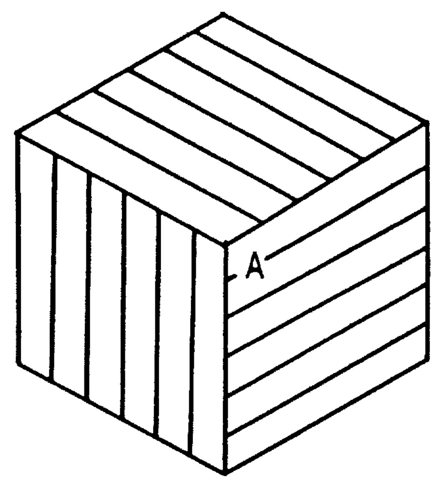


Fig. 2: Hinton's cube viewed from above

Why does the first version of the problem appear difficult while the second one seems easy? We may speculate that many people perceive a cube through a model based on Cartesian coordinates. In this framework, the angle of 90° (as well as multiples and parts thereof) are priviledged. The spatially diagonal axis and the rotation cannot be easily represented in this framework. In contrast, in a polar coordinate system, the angle of 360° (and parts thereof) are priviledged. The given rotation can be easily represented; the correct answer can be deduced on the basis of symmetry considerations. In this way, empirical psychological experiments can be used to build computational models and thus enrich both disciplines involved.

1.4 Knowledge and knowledge representation

The nature of knowledge is studied in epistemology, a subfield of philosophy. The field of logic has provided ways to formalize (mainly language-based) knowledge beginning in the last century with Frege. Artificial intelligence has built on the results of this research to implement knowledge-based systems. Linguists and psychologists have compared the behavior of artificial systems with the behavior of natural ones. A goal of cognitive science is to learn more about hidden knowledge and inference structures of cognitive systems, not only about their behavior. In particular, cognitive science is concerned with identifying suitable knowledge representation schemes for cognitive processes.

For the purposes of this paper, we can put 'knowledge = data + interpretation', where *data* can be anything and *interpretation* involves decoding and inferencing. 'Data' mirrors the static aspect, 'interpretation' the dynamic aspect of knowledge. The dynamic performance patterns are strongly dependent on the (static) data structures. The knowledge may reflect facts or beliefs about objects, actions, events, know-how, and metaknowledge.

This report reviews issues related to the representation of knowledge in cognitive systems. Representations are important for the study of knowledge, because they are the only way to describe knowledge structures. What we mean by a 'representation', will be defined in more detail in section 3. Much confusion and controversy has been created by the fact that there are not yet universally accepted concepts and a unified terminology in the field. With this in mind, we will attempt to clarify some of the terms and concepts under discussion within a unified framework.

1.5 Representation theory

We will promote the idea of discussing representations within a representation theory capturing objectives and properties of representations. A motivation behind this approach is that we would like to be able to construct adequate representation systems from task descriptions. Two main points suggested by a representation-theoretical approach to cognition are that 1) representations cannot be compared outside the framework of a representation system, and 2) the study of inference mechanisms should be preceded by the study of representation systems. Depending on the specific task to be performed, one representation system may be preferable over another (c.f. the paper by Luc Steels in this book). This raises questions about the transformation between various representation systems and about the computational cost associated with such transformations.

The present paper is related to work which has integrated results from psychology, computer science, philosophy, the neurosciences, logic, and linguistics exemplified by researchers like Zenon Pylyshyn, Stephen Palmer, Marvin Minsky, Aaron Sloman, David Marr, Pat Hayes, and Terry Winograd.

2 THE RELATION BETWEEN REPRESENTATION AND COGNITION

The problem of representing knowledge has been a central focus of artificial intelligence research from its beginning. In psychology, the significance of knowledge representation for understanding cognitive processes has been recognized in recent years as well. This is partly due to results from artificial intelligence, experimental cognitive psychology, and philosophical considerations.

Various foci of interest have determined research in knowledge representation: the search for *general* representations for problem solving; the search for *computationally efficient* representations; the search for *complete* and *consistent* representations for theorem proving; and the search for *natural* representations (i.e. occuring in nature) for cognitive modeling. The following sections review some representational issues in AI, psychology, and philosophy and the metaphors on which models of cognition have been based.

2.1 Representation of knowledge in artificial intelligence

In artificial intelligence research, a variety of approaches to knowledge representation have been explored, namely the logic approach, the semantic net approach, procedural representations, production systems, frame-based representations, and direct or analogical representations. Surveys of these approaches can be found in the *Handbook of Artificial Intelligence* (Barr & Feigenbaum 1981) and in the overview paper by Mylopoulos and Levesque (1983).

The different approaches to knowledge representation do not mutually exclude each other; rather they emphasize different ways of looking at the structures to be represented. The issues we are going to discuss in the present paper apply to various of these approaches to a greater or lesser extent.

The *logic approach*, for example, emphasizes (static) syntactic aspects of representation and the use of operators for inferencing. The (dynamic) control procedure transforming one logic formulation into another, is generally not described in terms of logic. During the transformation it may happen that certain knowledge states created do not have an interpretation in the represented domain.

The *semantic net approach* can be viewed as a notational variant of predicate calculus. However, its main focus is to emphasize the structural relations in the represented conceptual domain.

Procedural representations may be 'logically equivalent' to some declarative forms of representation. But this equivalence only describes the functional aspects of procedures. It does not capture the ways by which different states of knowledge are arrived at, i.e. its operational semantics.

Production systems are related to procedural systems in that they contain rules involving specific knowledge but give less emphasis to the process of applying the rules.

Frame-based approaches extend the structural features of the semantic net approach in that they provide for hierarchical organization of conceptual units and built-in procedural features.

Analogical (or *direct*) representations emphasize structural correspondence between objects, relations, and processes in the represented and the representing worlds.

2.2 The issue of representation in psychology

The present view of representation in psychology is most transparent from a historical perspective. From the 1930s to the mid-50s, the most influential school in psychology was *behaviorism*. Behaviorism restricts the scope of its theoretical concepts explicitly to the *performance level* of the investigated systems, i.e. humans or animals. Behaviorists did not attempt to open the "black box" to uncover control mechanisms like electrical and chemical processes on the physical (brain) level of description, and cognitive processes on the mental (mind) level of description.

Around the middle of the 1950s, however, some psychologists -- encouraged by a growing understanding of information, computation, and 'automatic thinking machines' -- tried to open the black box (c.f. Neisser 1967). It had been clear since the beginning of scientific psychology that stimulus-response chains in general were based on the utilization of knowledge. In particular, the question had been asked *what* knowledge was required. The importance of *how* this knowledge was represented remained to be discovered later.

The *cognitive turn* in American psychology was marked by a significant extension of the scope of theoretical concepts for describing the contents of the black box and by the acceptance of the *information processing metaphor* (Miller 1956; c.f. section 2.4) from computer science as the dominant framework for theorizing. After these developments, in particular after becoming familiar with the information processing metaphor the need for concepts for knowledge representation was discovered by psychologists. In the meantime, in computer science and AI attractive representation schemes had been suggested. The experimental work in psychology stimulated by this paradigm shift was mainly aimed at demonstrating the compatibility of human cognitive phenomena with models developed within the information processing framework.

There are three classical *windows to cognition* through which psychologists can peek into the black box, three types of observational variables which may tell us something about the functional structure of the ongoing processes: 1) frequency counts of reactions, 2) measures of response times, and 3) the comparison of 'correct' with 'erroneous' reactions. Equipped with this "toolkit", psychologists have invented numerous experimental paradigms to study the representation of knowledge in memory. The objective of the experimental studies was to demonstrate congruences of observable effects with the predictions of models for representation. All major representational schemes have been explored in this way. We will demonstrate the flavor of this work by two examples, namely semantic networks and analogical representations.

Semantic networks were understood to be suitable for capturing the semantic relations between concepts. Subjects were asked to judge propositions about concepts as to whether they were true or false ("Is a penguin a bird?"). Measurements of the response time were interpreted as indicators for the graph structure of the activated knowledge: longer response times indicated that a greater number of links had to be followed up until the answer was reached (c.f. Collins & Loftus 1975).

Analogical representations (c.f. section 4.3) seemed quite plausible as a form of representation in memory. However, they do not emerge in as natural a way within the information processing framework as do semantic networks, for example. Evidence for their availability as representational form in memory was provided from experiments on the 'mental rotation' of visually presented figures, for example drawings of 3-dimensional block figures. The subjects were asked to judge whether two adjacent figures shown in different positions were identical or not. The finding was that response times were proportional to the spatial angle of rotation between the two figures (Shepard & Cooper 1982).

The controversies and debates about the issue of representation in AI were echoed by controversies and debates in psychology and vice versa. A full understanding of these issues requires looking at both, the developments in AI and the simultaneous stream of work in psychology.

2.3 Philosophical considerations

Issues related to knowledge and cognitive representation have concerned philosophers for a long time. George Berkeley distinguishes in his *Essay towards a new theory of vision* (1709) between objects and their representations and in his *Principles of human knowledge* (1710) he maintains that 'physical objects can exist only in the mind' (Berkeley 1965).

In his *Begriffsschrift*, Gotlieb Frege (1879) investigates the identity relation and argues that it is a relation between names or symbols rather than a relation between objects. In his paper *über Sinn und Bedeutung* Frege (1892) discusses the difference between Sinn (sense or meaning) and Bedeutung (denotation) of representations.

Bertrand Russell (1905) points out the problems with constructing a theory of denotation and discusses the issue of representing false facts. He develops a theory of knowledge in which he advocates that there is no meaning in symbols or denoting phrases but only in the propositions in which they occur.

John McCarthy and Pat Hayes (1969) point out the importance of philosophical considerations for an advancement in artificial intelligence. They argue that computer programs require a metaphysically and epistemologically adequate general world representation in order to interpret their inputs intelligently and formalize concepts of situation, action, strategy, result of a strategy and knowledge, etc. which they consider basic elements of intellectual mechanisms.

Aaron Sloman (1971) also emphasizes the importance of relating philosophical issues to the design of intelligent machines but criticizes McCarthy and Hayes for only considering languages like predicate calculus and programming languages for the formalization of concepts, and not, for instance, the 'language' of maps. He aims at generalizing the concept of a valid inference to include non-verbal representations as a basis for rigorous reasoning. Sloman takes a strong position towards directly representing spatio-temporal environments in order to allow for efficient analogical reasoning rather than describing everything in terms of general Fregean logic representations. In his follow-up paper, Sloman (1975) precisiates his arguments in the light of a more intimate knowledge about computers. He acknowledges that AI programs had been using analogical representations for efficiency, but this fact had not been made explicit.

Hubert Dreyfus (1979) believes that intelligence is intimately dependent on the physical nature of human beings and their sensory-motor system. He denies that intelligent behavior can be imitated within an artificial medium like a computer which does not share the properties he considers essential.

John Searle (1980) makes a distinction between two positions supposedly held by AI researchers, the *weak* and the *strong* AI. While the weak position merely employs the computer as a powerful tool, the strong position maintains that appropriately programmed computers have cognitive states, and therefore the programs are psychological theories. According to Searle, this strong AI position must be false.

2.4 The information processing framework and the notion of a symbol system

We indicated in the introduction that it is not easy to define the term *cognition*. The first attempts to investigate human cognition in psychology were studies on remembering and forgetting. The common metaphor then was the concept of *association* (Baddeley 1976). It turned out that many cognitive phenomena could not be exthe interest and expectations of psychologists in the *information processing paradigm* introduced by Miller (1956) and Simon and Newell (1956).

The information processing paradigm may be the best approximation to a definition of cognition -- cognitive processes are those processes in the brain or in artificial devices which generate or transform information. This approximate identification of cognition with information processing has determined theorizing in cognitive psychology and AI since the late 1950s.

Today's digital computers are *von Neumann machines* which are closely related to one another with respect to their computational primitives. They can be described as *symbol processing machines* in a natural way. Alan Newell (1980) equated the mind with a *physical symbol system*. On one hand, this precisiation of the information processing metaphor was accepted, on the other hand, the possibility of exploring computational devices of a different kind was provoked (c.f. Hinton & Anderson 1981).

The central claim of Newell is that the functional structure of the black box, i.e. the operation of the mind, can be precisely described in terms of a symbol system. Symbol systems are Turing machines. According to this theory, the mind is a system capable of generating and transforming symbols. Cognitive processes are symbol manipulating processes. Thus, the theory of symbol processing is viewed as an appropriate theory of cognition.

The symbol system metaphor is a precisiation of the information processing metaphor. There are two consequences of this theoretical refinement that should be noted here. First, the symbol system provides a sound theoretical foundation for knowledge representation and the different representational schemes, a fact which is obvious to computer scientists but has been less evident to psychologists. Second, the symbol system framework restricts the spectrum of information processing systems, for example an optical microscope can be viewed as an information processing device, but not as a symbol processing system.

A possible source of confusion in connection with the symbol system metaphor must be ruled out. No one will seriously question that many cognitive processes are concerned with symbols. For example doing a mathematical derivation of a new theorem or communicating by language involves symbol manipulations. However, these symbols must be distinguished from the symbols processed inside the 'black box' of the mind. For the former symbols denote entities in the external world like mathematical objects in the abstract world of mathematics or objects like trees or dogs in the "real" world, whereas the latter symbols only point to other symbols inside the mind.

3 THE NOTION OF A REPRESENTATION SYSTEM

In computer science, symbol systems have been studied from the very beginning and the problem of representation is well understood. The acceptance of the equation *cognition = symbol system* emphasized the importance of the issue of representation in cognitive science (Palmer 1978). In the following, we shall discuss the problem of representation from a perspective which may be suitable for symbol systems as well as for other computational approaches to cognition. As a starting point we chose Steve Palmer's notion of a *representation system*.

Palmer, a cognitive psychologist, is concerned with human cognition, i.e., with the representation of structures of the real world in the human mind. In our work, we want to understand in addition the representation in intelligent artificial systems from a more formal representation-theoretical point of view. In order to do this, we have to study two kinds of representation tasks: 1) representation of structures from the external world in artificial systems and 2) representation of structures entirely within such systems.

According to Palmer, a representation system basically consists of two worlds, a *represented world* and a *representing world*, which are related to one another by a correspondence mapping. In order to specify this correspondence, it is necessary to state which aspects of the represented world are to be modelled and which aspects of the representing world are doing the modelling. In order to characterize the notion of an "aspect of a world" we should note that in the information processing framework it is fundamental that a discussion of representation is done on the basis of processes. Thus, we regard relations between objects of a world as given operationally by processes.

For example, assume a world of blocks and a robot with a visual perception system capable of discriminating between blocks and of performing actions in this world. This robot -- regarded as a process -- defines relations in this world. For example, it can find two blocks and by interpreting their height it can define the relation *taller-than*.

Palmer (1978) included relation-defining processes only implicitly in his notion of a representation system. For his purposes this is sufficient, since he aims at modelling mental representations of the real world and can thus assume that 'we all have more or less the same operational concepts about relations in the world'. In this case, we can specify sufficiently precisely which aspects of a world are contained in a representation system. If, however, a representation system is to be used in the context of artificial systems, the notion requires further refinement.

If we were to design a knowledge representation system for our robot's environtures and procedures). For the represented world of blocks, we can use our concepts about blocks and about relations, for the representing world (the robot's memory), we have to define data structures and programs to work on them.

We include these object and structure defining relations explicitly in our notion of a representation system by associating with both worlds -- the represented as well as the representing world -- a so-called *reference world* which contains these relations (Furbach et al. 1984). These reference worlds can be regarded as interpreters of the worlds they are associated with. We will call a pair consisting of a world and its associated reference world, a *body of knowledge*. The correspondence between a represented and a representing world is given by a mapping from one body of knowledge to another.

We distinguish five components of a knowledge representation system (KRS):

1) a *represented world* W_1,
2) its associated *reference world* R_1,
3) a *representing world* W_2,
4) its associated *reference world* R_2, and
5) a *mapping* C establishing the correspondence between the represented body of knowledge (W_1, R_1) and the representing body of knowledge (W_2, R_2).

In the example given above, there is one body of knowledge consisting of the world of blocks (the represented world) and its associated reference world. In the case of *real world objects*, the reference world can be characterized only by naming our concepts about relations in this world. The second body of knowledge, however, can be defined completely by specifying the representing world (the robot's memory structure), and its associated reference world (the robot's programs and control devices). These two bodies of knowledge can be referred to by the correspondence mapping.

The separation between worlds and reference worlds is suggested by the computer metaphor. In computers we are able to distinguish easily between the world per se (or the data) and the artifact dealing with the world (or the program). This separation is less obvious in natural cognitive systems where the borders are defined less clearly.

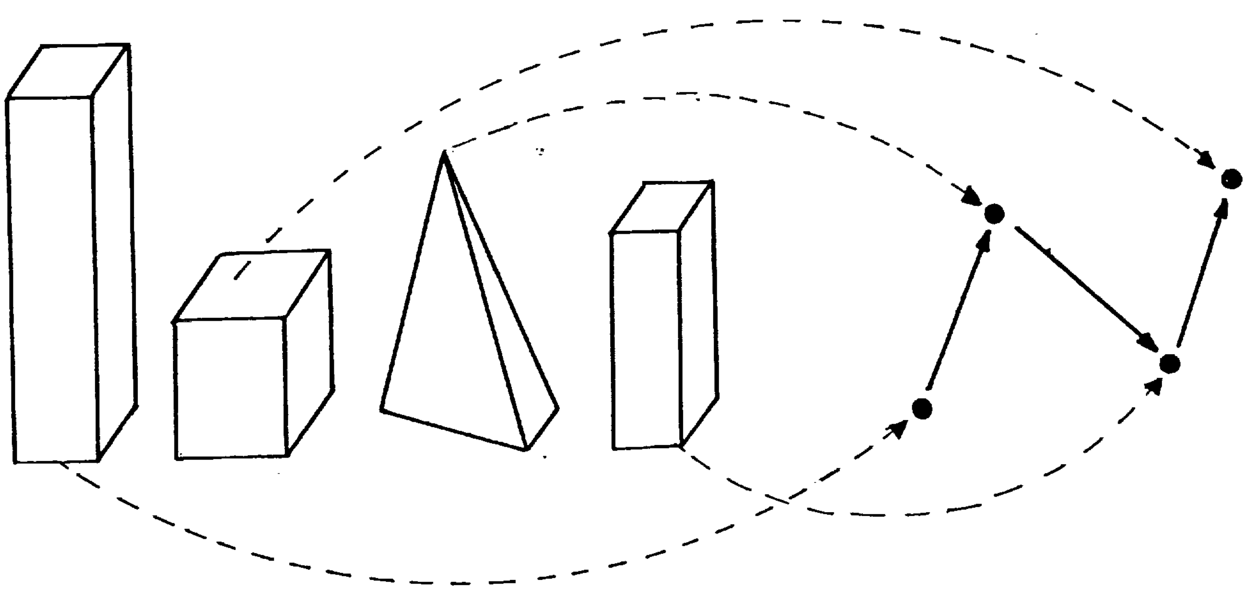
In contrast to the rich literature on knowledge representation formalisms we do not focus on a specific formalized language for defining worlds and reference worlds. Instead, we want to point out that it is necessary to look at complete representation systems, i.e. two bodies of knowledge together with their correspondence mapping. We will demonstrate that this view of knowledge representation enables us to be specific about the aspects of a world which are considered relevant in a given representation

3.1 Intrinsic and extrinsic representations of world aspects

In order to talk about properties of representation systems we introduce the toy representation system depicted in Fig. 3. The body of knowledge (W_1, R_1) consists of a simple world of blocks W_1 (a robot's environment) together with one object and one structure defining relation in the reference world R_1. The representing body of knowledge (W_2, R_2) (the robot's data strucures and procedures) consists of a world W_2 in which two kinds of objects are defined by R_2, namely dots and arrows, together with the structure-defining relation *chains-to*. The relation *taller-than* of W_1 is modeled by the relation *chains-to*.

The relation *taller-than* in R_1, as well as the corresponding relation *chains-to* in R_2 hold for all six pairs of vertical lines in W_1 and for the dots in W_2, respectively. *taller-than* is a transitive relation; thus we may expect the corresponding relation to have the same property. There are representing worlds which provide relations that share the transitivity property with *taller-than*. *heavier-than* would be such a relation, and *chains-to* is also one of them. This is an *intrinsic* form of representing properties of relations. 'Intrinsic' means, that a given property (e.g. transitivity) is inherent both in the represented relation and in the corresponding representing relation.

The relation *taller-than* has a second inherent property, namely asymmetry. *heavier-than* shares this property with *taller-than* but *chains-to* does not. The relation *chains-to* is not inherently asymmetric, this property is only maintained by excluding symmetric chains. This is an *extrinsic* form of representation.



is-object (block) -----------> is-object (dot)
is-object (arrow)

taller-than (object$_1$, object$_2$) -----------> chains-to (dot$_1$, (arrows), dot$_2$)

Note that it is easy to modify the system above such that transitivity of *taller-than* is represented extrinsically as well. We only have to replace the relation *chains-to* in R_2 by a relation *arrow-connected* and place additional arrows in W_2, such that we obtain the transitive closure of the graph W_2.

The examples above show that even in representing a single relation both modes of representation can be used simultaneously: one property can be represented extrinsically, while another can be represented intrinsically. In section 4.3 we shall use these different ways of representing properties of relations to characterize the controversial issue of analogical representations.

Another example particularly relevant to physical environments presents the *uniqueness* property of locations of physical objects in space. If we represent physical objects by copyable data in a computer, the inherent uniqueness of the object location relation is not preserved and we must ensure extrinsically that this property is mimicked. This can be done, for example, by declaring the state of a representation system 'undefined' when a data object is being copied and the original instance has not yet been eliminated.

In general, intrinsic representations of aspects are less complex and involve less complex operations than extrinsic ones since necessary properties do not have to be explicitly specified. However, most reasoning operations only can be performed on explicit representations making certain aspects extrinsic. A guideline for selecting a certain representation scheme therefore is to keep intrinsic as many properties as possible and to make extrinsic only the ones that are needed for the reasoning process.

3.2 The boundary between a world and its reference world

We conceive the reference world as a set of processes which are accessible for investigation of their properties with respect to the given representation system. Processes in the reference world merely define the structure of the world to which they are assigned; however, they can not modify this world as such.

In contrast to this situation, we may be interested in the *results* generated by processes of the reference world: we may want to add these results as new objects to the world. Furthermore, we may be interested in treating processes as objects in the world. In this case, we must shift them from the *reference world* into the *world*. Such process-objects in the world would not be open for "inspection" any more. They could, however, still be triggered and produce results. We will use the distinction between *world* and *reference world* to discuss the distinction between procedural and

4 PROPERTIES OF REPRESENTATIONS

After defining the framework of representation systems, we shall focus our attention on properties by which various systems have been classified. We will talk about *iconic, analog, analogical, propositional, procedural, declarative, Fregean, logic, linguistic,* or *non-linguistic* representations. These concepts and the debates about them (c.f. Sloman 1971; Hayes 1974; Winograd 1975; Pylyshyn 1981; Brachman 1983) can only be fully understood within their historical contexts. They emerged during the development of information processing approaches within the various disciplines of cognitive science. In most cases, the debates were concerned with specific aspects of representations rather than with complete representation systems, a fact often forgotten or ignored and a cause of confusion in debates on knowledge representation.

Note that the concepts and controversies are rooted in the common acceptance of the *symbol system metaphor* as a theory of cognition. We shall demonstrate in the following sections that within our generalized notion of a representation system some of the issues under discussion appear less controversial. Moreover, we shall show that the concepts developed can be used for the design of representation systems emphasizing problem-specific representation schemes.

4.1 Iconic representations

A representation consists of objects and relations between them. *Iconic* forms of representation are representations in which certain (or all) spatial relations are preserved, i.e., these relations are identical in the represented and the representing worlds. An example of a representation in which all spatial relations are preserved is a photograph of a 2-dimensional object taken perpendicular to the object plane. A photograph taken from a different angle or a photograph of a 3-dimensional scene only preserves some of the spatial relations. Strictly speaking, such representations should be called iconic only with respect to the preserved spatial relations.

Iconic representations allow the same relations for interpretation of spatial aspects to be used in the reference world of the representing world as in the reference world of the represented world. As a consequence, spatial reasoning processes may be carried out in much the same way in the representing world as spatial operations in the represented world. This aspect has contributed much to the debate on *mental imagery* (c.f. Anderson 1978; Block 1981; Pylyshyn 1973, 1981, 1984; Kosslyn 1975; Kosslyn et al. 1978) which has been one of the core topics in cognitive science during the past decade.

eye'). This way of representing, therefore, would not explain the process of scene understanding. It merely would defer the problem by one level, eventually leading to an infinite regress".

Proponents of iconic representations point at empirical evidence relating perceptual response times of humans to spatial aspects of the perceived objects (e.g. Shepard's *mental rotation* experiments outlined in section 2.2). The infinite regress thesis has been countered with the argument that a first iconic representation preserving spatial properties can be re-represented for specific tasks (*dual code theory*). This is an interesting idea from a computer science point of view and has been explored by David Marr (1982) in his *primal sketch theory*. The dual code theory appears attractive for representations in intelligent systems since one code preserves spatial aspects. This implies that the representation for these aspects is intrinsic. Task-oriented transformations of this representation can be built upon these 'primal sketches', making extrinsic only those properties which are needed for the given task (Dirlich et al 1983).

In cognitive science, two fundamental forms of knowledge representation have been discerned, *mental images* and *propositional representations* (Anderson 1978). Mental images are iconic representations and have been regarded contradictory to symbolic mental representations. From our KRS-perspective, we distinguish between *relation-preserving* and *structure-preserving* representation systems. According to this distinction, iconic representations are relation-preserving. However, experimental investigation of the performance of representation systems can only indicate something about the represented structure, but not about the relations which generate the structure. In the example of human visual capabilities we therefore should not infer the existence of mental images in the literal sense but only the existence of some representation structure which behaves like an image (c.f. Shepard & Cooper 1982). Such representations are discussed below.

4.2 Analog and digital representations

For historical reasons, the term *analog* has been used in the computer literature specifically to distinguish continuous representations of continuous features from inherently discrete representations which are called *digital* (c.f. Pylyshyn 1984; Brachman and Smith 1980, p.87). This is a very restricted special case of analogy (namely with respect to the continuity property). It becomes meaningless if the represented world is discrete, since an 'analog' representation of a discrete world does not preserve the property 'discrete' in an analogous way.

Analog versus *digital* has created some debate in artificial intelligence. Notably, Hubert Dreyfus (1979) maintains that analog transmission and handling of in-the argument that much of the information between nerve cells is transmitted digitally by frequency coding. (Digital-coded information can be transmitted more reliably than analog-coded information, as hi-fi fans know.)

The analog-digital distinction may be viewed as a change in the level of description. Digital signal systems like nerve structures and digital computers are based on analog phenomena, the electric currents in nerve cells and chips. These, in turn, are based on discrete units, the elementary charges of electrons. Thus, *analog* and *digital* may refer to different levels of description of the same system. Digits make up the material out of which symbols are built (in digital computers). Thus, *analog - digital* can be related to the *physical process - symbol process* distinction of representations.

4.3 Analogical representations

In cognitive science, *analogical* (rather than 'analog') representations have created more of an issue. The term 'analogical' can refer to a multitude of aspects in representation systems; it means that some aspect in W_2 (under the interpretation of R_2) is structured the same way as the corresponding aspect in W_1 (under the interpretation of R_1). The term 'analogical representation' has been contrasted to 'linguistic', 'Fregean' and 'logic' representation and has little to do with the analog - digital distinction.

Sloman (1975) clarifies a number of misconceptions about analogical representations. Specifically, he points out and exemplifies that analogical representations do not have to coincide with continuous, multi-dimensional, isomorphic, non-symbolic, complete, or non-grammatical representations. Rather, 'analogical' refers to the structural correspondence between certain aspects in the representing and represented worlds whereas in the case of Frégean representations there need be no structural correspondence. In particular, aspects represented in an analogical way will agree in complexity in the representing and represented worlds.

Sloman (in Brachman & Smith 1980, p.128) argues that for an understanding of intelligence it is important '*how* things are represented, not *what* is represented'. This is why some workers in cognitive science are specifically interested in analogical representations. For an elaborate discussion of the difference between analogical and Fregean representations, see Sloman (1975, p.179).

To avoid confusion, only representation systems in which all represented aspects (as specified by correspondence mapping C) are represented analogically should be called 'analogical representation systems'. Otherwise, the analogically represented aspects should be clearly stated.

elements are ordered according to the order of what they represent. In this case the relation *next-to* is used to represent the relation *is-neighbor* in the represented body of knowledge. The properties irreflexivity and asymmetry, for example, are inherent in both relations, i.e. these properties are represented intrinsically.

In the framework of a KRS we can identify analogical representations with intrinsic representations of properties. As a consequence, we can restate more precisely that the term 'analogical', (or 'intrinsic') refers to the way properties of relations are represented in a given representation system. Much of the confusion about analog and analogical representations (c.f. Pylyshyn 1984) could be avoided if it was realized that *analogy* relates two structures to one another and does not refer to an absolute feature.

4.4 Propositional representations

Most of the representation systems developed in AI are propositional systems. Under the term *propositional* we include both declarative and procedural types of representations; logic representations can be regarded as prototypical.

In philosophy and in formal logic a proposition is an entity which can be true or false. In logic systems, propositions are symbols which can be interpreted only by the two alternative truth values. In a strong sense, such an interpretation is not an assumption about reality, it only serves as a basis for investigating relations about these propositions. For this purpose, logic systems provide for the possibility of constructing more complex formulas and of deriving new formulas by means of inference mechanisms. This can be of great use for the issue of knowledge representation because one is forced to make each significant aspect of a represented world individually explicit. In this way it is possible to discover inconsistencies and to make logical inferences over the entire range of represented knowledge by means of a uniform 'inference engine'.

On the other hand, just this property caused many workers to consider alternative forms of representation. First, pure logic formalisms provide few tools for structuring knowledge in memory. This led to the exploration of representational forms which emphasize more psychological aspects, namely

- the notion of knowledge 'units', so that knowledge about single concepts or events is organized according to functional units,
- the detailed structure of knowledge about single concepts or events, and
- the consideration of different levels of kowledge.

Second, the inference mechanisms of logic systems, though powerful enough for a real-
ization within artificial systems, do not capture the ways people seem to reason. In

explicitly, they apply information of one concept to another and they use inconsistent knowledge. Attempts to incorporate such forms of reasoning into logic systems have resulted in structures that are rarely considered 'natural'.

Logic-based formalisms can be enriched with components from other forms of representations. For example, the 'connection method' for theorem proving (Bibel 1982) uses graphical constructs (pointers) to reduce redundancy that is found in pure logic formalisms. Pure logic representations are characterized by the property that all representational aspects to be considered by an interpreter or inference process are made explicit by symbolic expressions in a uniform way. In other words, in pure logic representations all properties of relations from within the represented body of knowledge are represented extrinsically -- the relations of the representing (i.e. the logic) body of knowledge have no inherent properties.

Another disadvantage of a uniformly structured, purely syntactic approach is a potential increase in processing complexity. During an inference process, many rules may be applicable for purely syntactic reasons without meaningful semantic interpretation. As a consequence, large problem spaces may be created by combinatorial explosion. In essence, this is also due to the fact that only extrinsic representations are used.

4.5 Declarative and procedural representations

Consider three information processing situations, 1) the process of answering the question 'what kind of an animal is a robin?', 2) the process of answering 'what is 306 divided by 18?', and 3) the process of responding to a visually perceived approaching tennis ball. In the three situations, apparently different types of knowledge structures are activated. In the first case we may associate a retrieval procedure in a data bank, in the second case we may think of a computational procedure, and in the third case a highly complex cybernetic system seems to be triggered. The differing knowledge structures utilized in these cases also differ in another respect: some of them seem to be open to conscious inspection whereas others are not.

To understand these different types of knowledge by our notion of a KRS, assume we have the task of building the basic LISP list-processing primitives in a programming language like PL/I. We then can use data types like *pointers* and machine-level *address operators*. The actual operations of the program are here transparent. We may call this way of representing knowledge *declarative*.

After having defined PL/I procedures for *car*, *cdr*, and *cons*, we may want to disallow the use of *pointers* and *address operators*, we only allow the use of the list vated to yield results. We may call this hidden way of representing knowledge *procedural*. In computer science the approach of shifting a process from the declarative to the procedural level is known by the name *data abstraction*.

In the KRS framework, hiding the internal structure of a relation defining process corresponds to the shift of a process from the reference world into the world. This means, the boundary between a world and its reference world (c.f. section 3.2) within a body of knowledge separates declarative from procedural knowledge. Relations within a world only can be activated to yield a result, they define procedural knowledge. Relations within a reference world are open for inspection, they define declarative knowledge.

Rumelhart and Norman (1983) describe a similar approach to this aspect of representation systems. They define a representation system RS as a *relational double*, $RS = \langle R, P \rangle$, where R is the representing world and P is the set of processes that operate upon and interpret R. Rumelhart and Norman use a different meaning for 'declarative' and 'procedural'. Motivated by considerations in psychology, they call R the 'declarative part' of the knowledge and P the 'procedural part'. With our KRS notion we arrive at a different procedural/declarative dichotomy. This is an example of the importance of making the reference system explicit when discussing representational issues (c.f. Winograd 1975).

5 PERSPECTIVES IN KNOWLEDGE REPRESENTATION THEORY

There is no theory of knowledge representation
Handbook of Artificial Intelligence

In the preceding chapters various aspects of knowledge representation have been discussed from the perspectives of psychology and AI. This discussion has shown that the field is still in a pretheoretic stage of development, mainly concerned with the invention and exploration of forms of representation by AI researchers. In psychology, interesting work has been done in order to map empirical knowledge onto the various aspects of representation systems discussed in section 4 (c.f. Norman & Rumelhart 1975; Anderson 1983).

Although we are aware of this situation, we are proposing a broader, more theoretical approach to knowledge representation. The concept of knowledge representation systems presented in section 3 can be viewed as a step in this direction. The objective of knowledge representation theory (KRT) could be captured by the following formula: "KRT should describe how appropriate task-specific representations can be

In the present concluding section we will mainly discuss possible benefits and problems of a theory with respect to knowledge engineering. In particular, we will discriminate between realistic goals that are likely to be reached in the near future and more speculative goals that nevertheless may stimulate scientific creativity.

5.1 Goals of knowledge representation theory

We showed in section 4 how KRT can contribute to a resolution of certain problems and controversies that stimulated a lot of scientific activity in recent years. The knowledge engineering task typically has two aspects, namely the aspect of the system in which the representation is done and the aspect of the knowledge that the representation is supposed to capture. KRT should support both aspects. Concretely, we can discern three types of tasks in knowledge engineering which might profit from the theory. Ordered by increasing difficulty of being tackled by KRT, they are

- the unified description of the properties of forms of knowledge representation,
- the description of relations between different bodies of knowledge,
- the conception of selection criteria for forms of representation.

The first task can be viewed from both, the system and the knowledge perspectives. From the system-centered perspective, knowledge representation formalisms first of all must be considered with respect to their cost. Even simple examples like the one concerning list processing (discussed in section 4.5) show how certain system-centered properties of forms of representation can be treated within a framework of representation theory. From the knowledge-centered perspective, properties appear to be still not well enough understood to become tractable within the theory. In production systems, for example, rules are based on a directed relation between the condition and the action component. This implies that the knowledge to be represented contains the aspect of directionality. In some instances, this may be the case and the form of representation perfectly fits the nature of the knowledge. In other instances, however, directionality may be too strong a relation, and 'co-occurence' of the condition and action components may provide a better form for the representation. How, do we analyze systematically, what may be good candidates for the form of representation for given but not yet fully specified tasks? Today, the task of selecting a suitable form of representation is usually intuitively solved by the knowledge engineer. A future KRT should provide a rationale for the selection.

Similar considerations can be carried out for the second type of task, the description of the relations between corresponding bodies of knowledge in multiple code representation systems. Here also, cost is a relevant aspect. The transition from one form of representation to another may significantly affect the 'difficulty' of inferencing. Obviously, there is a trade-off between cost for inferencing and cost to be a great scientific challenge to approach the decision problem concerning an appropriate or even good (optimal) form of representation by AI methods. Could representation theory provide criteria for a task related evaluation of candidate forms of representation?

Progress in this direction may eventually lead to a paradigm shift in AI. Instead of measuring the reasoning power of a system only by its power of inferencing, intelligent systems may contain powerful components for knowledge transformation on which their reasoning power may crucially depend. A future knowledge representation theory would impact knowledge engineering and in particular contribute to the development of design principles for human-machine interaction.

5.2 Formal versus illustrative representations

Most of us think in diagrams of one kind or another, and it is sometimes useful to others if we make these private visions public.

Jerry Hobbs (1980)

The author of the lines above states in the same paragraph: "Learn from mathematics the proper relation between diagrams and formalism. Diagrams are for illustration, not for formalization." This sounds like a contradiction: diagrams are useful for thinking and communicating, but they are something 'personal' and should not be used for serious reasoning.

What is the formal difference between propositional and graphical formalisms that makes the former generally accepted as a scientific tool for reasoning, while the latter is only accepted as an insufficient substitute. In the well-defined. domain of constructive geometry, graphical proofs are considered equivalent to symbolic ones which were invented later (c.f. Gelernter 1963). Sloman (1975) indicates that it may be a very difficult task to do formal reasoning on diagrams, maps, or other spatial structures. We would like to suggest, however, that by studying properties of representation systems, the domain of well-defined non-symbolic structures might be expanded in such a way that formal reasoning can be performed on illustrative representations other than geometrical ones, as well (c.f. Funt 1980, 1983).

5.3 Problems with knowledge representation theory

Several serious problems may be expected in the course of developing a theory of knowledge representation. We will mention three of them here.

Computer science and artificial intelligence are scientific domains which are rich in theory. Scientific progress, however, does not solely depend on the amount (e.g. development of human-machine interfaces using menues, window techniques, mouse, etc.). Too much theory may hide the need of working on the content of the scientific domain (Minsky 1970).

It has been demonstrated in the preceding sections that work in the area of knowledge representation is dominated by conceiving new forms of representation and exploring their appropriateness for different task domains. Any theory must be build upon a suffiently broad basis of experience with the objects in the domain. In the present case it is not evident if this basis is already sufficient. If it is not it would be bad timing to engage now into attempts to develop knowledge representation theory. However, we want to point out here that we are aware of a gap that should be filled by knowledge representation theory.

An area of science which is in some respect adjacent to the area of knowledge representation is the theory of measurement which is at the foundation of oberservation and data analysis in the empirical sciences, where soft data play an important role. Measurement theory is aimed at the mathematical description of the relation between the to-be-measured phenomena in the real world and the abstract objects which are the result of the measurement. The objective of measurement theory is to prove theorems about these relations under certain constraints in the real world and in the representing data world. So far, the existence of relations that fulfill the given constraints has been proven only for simple cases (e.g. Luce et al. 1963).

Acknowledgements

Part of the work reported in this paper was done while one of the authors (C.F.) was visiting Lotfi Zadeh's stimulating research environment at the University of California, Berkeley. Wolfgang Bibel's efforts which made this project possible are much appreciated. Steve Palmer and Paul Smolensky provided a lot of food for thought. Dennis Allard made critical comments on drafts of this paper.

REFERENCES AND SELECTED BIBLIOGRAPHY

ANDERSON, J.A., HINTON, G.E. 1981. Models of information processing in the brain. In: Hinton, G.E., Anderson, J.A. (1981).

ANDERSON, J.R. 1978. Argument concerning representations for mental imagery. Psych. Review 85, 249-277.

ANDERSON, J.R. 1983. The architecture of cognition. Harvard, Cambridge, Mass.

BADDELEY, A.D. 1976. The psychology of memory. Basic Books, New York.

BARR, A., FEIGENBAUM, E. 1981. The handbook of artificial intelligence. Vol. 1, Heuristech Press, Stanford.

BIBEL, W. 1982. Automated theorem proving. Vieweg, Braunschweig.

BLACK, M. 1972. How do pictures represent? In: Art, perception and reality, E.H. Gombrich, J. Hochberg, M. Black (eds.), Baltimore.

BLOCK, N.J., ed. 1981. Imagery. MIT Press, Cambridge, Mass.

BOBROW, D.G., COLLINS, A. 1975. Representation and understanding. Academic Press, New York.

BRACHMAN, R.J. 1983. What isa is and isn't: an analysis of taxonomic links in semantic networks. IEEE Computer 16, 10 (Special issue on knowledge representation).

BRACHMAN, R.J., SMITH, B.C, eds. 1980. Special issue on knowledge representation. SIGART Newsletter No. 70.

CHAFE, W.L. 1975. Creativity in verbalization as evidence for analogic knowledge. TINLAP-1, 158-159.

CLOWES, M.B. 1971. On seeing things. Artificial Intelligence 2, 79-116.

COLLINS, A.M., LOFTUS, E.F. 1975. A spreading activation theory of semantic processing. Psych. Review 82, 407-428.

DIRLICH, G., FREKSA, C., FURBACH, U. 1983. A central problem in representing human knowledge in artificial systems: the transformation of intrinsic into extrinsic representations. Proc. 5th Cognitive Science Conference, Rochester, New York.

DREYFUS, H.L. 1979. What computers can't do: a critique of artificial reason. 2nd ed. Harper & Row, New York.

EVANS, T.G. 1968. A program for the solution of a class of geometric-analogy intelligence-test questions. In: Minsky (1968).

FEIGENBAUM, E.A., FELDMAN, J. 1963. Computers and thought. McGraw Hill.

FREGE, G. 1879. Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. L. Nebert, Halle.

FREGE, G. 1892. über Sinn und Bedeutung. Zeitschrift für Philosophie und philosophische Kritik, vol. 100, 25-50, Leipzig.

FUNT, B.V. 1980. Problem-solving with diagrammatic representations. Artificial Intelligence 13, 201-230.

FUNT, B.V. 1983. Analogical modes of reasoning and process modeling. IEEE Computer 16, 10, 99-105.

FURBACH, U., DIRLICH, G., FREKSA, C. 1984. Towards a theory of knowledge representation systems. To appear in: Bibel, W. & Petkoff, B. (eds.) Artificial intelligence methodology systems applications - AIMSA '84. North-Holland, Amsterdam.

GELERNTER, H. 1963. Realisation of a geometry-theorem proving machine. In: Feigenbaum & Feldman (1963), 153-163.

HAUGELAND, J., ed. 1981. Mind design. MIT Press, Cambridge, Mass.

HINTON, G.E. 1980. Some demonstrations of the effects of structural descriptions in mental imagery. University of California, San Diego.

HINTON, G.E., ANDERSON, J.A., eds. 1981. Parallel models of associative memory. Erlbaum, Hillsdale, N.J.

HOBBS, J. 1980. In: Brachman & Smith (1980), p.44.

HUNT, M. 1982. The universe within -- a new science explores the human mind. Simon and Schuster, New York.

JOHNSON-LAIRD, P.N., WASON, P.C., eds. 1977. Thinking. Cambridge University Press.

JOHNSON-LAIRD, P.N. 1980. Mental models in cognitive science. Cognitive Science 4, 1, 71-115.

KAY, P. 1981. Color perception and the meaning of color words. University of California, Berkeley.

KOSSLYN, S.M. 1975. Information representation in visual images. Cognitive Psychology 7, 341-370.

KOSSLYN, S.M. 1980. Image and mind. Harvard University Press.

KOSSLYN, S.M., BALL, T.M., REISER, B.J. 1978. Visual images preserve metric spatial information: evidence from studies of image scanning. J. Exp. Psych.: Human perception and performance 4, 46-60.

KOWALSKI, R. 1974. Predicate logic as a programming language. IFIP-74, 569-574, North-Holland, Amsterdam.

LUCE, R.D., BUSH, R.R., GALANTER, E., eds. 1963. Handbook of mathematical psychology. Wiley, New York.

MARR, D. 1982. Vision. Freeman, San Francisco.

McCARTHY, J., HAYES, P. 1969. Some philosophical problems from the standpoint of artificial intelligence. Machine Intelligence 4, 463-502.

MILLER, G.A. 1956. The magical number seven, plus or minus two. Psych. Review 63, 81-97.

MINSKY, M.L. 1970. Form und content in computer science. ACM Turing lecture. JACM 17, 2, 197-215, 1970.

MINSKY, M.L. 1975. A framework for representing knowledge. In: Winston, P. (ed.) The psychology of computer vision. McGraw Hill, New York.

MYLOPOULOS, J., LEVESQUE, H. 1983. An overview of knowledge representation. GWAI-83, 143-157. Springer, Heidelberg.

NEISSER, U. 1967. Cognitive Psychology. Appleton-Century-Crofts, New York.

NEWELL, A. 1980. Physical symbol systems. Cognitive Science 4, 2, 135-183.

NEWELL, A., SIMON, H. 1972. Human problem solving. Prentice Hall.

NORMAN, D.A., RUMELHART, D.E., eds. 1975. Explorations in cognition. Freeman, San Francisco.

PALMER, S.E. 1975. The nature of perceptual representation: an examination of the

PALMER, S.E. 1978. Fundamental aspects of cognitive representation. In: Rosch, E. & Lloyd, B.B. (eds.) Cognition and categorization. Lawrence Erlbaum, Hillsdale.

PYLYSHYN, Z. 1973. What the mind's eye tells the mind's brain: a critique of mental imagery. Psych. Bull. 80, 1, 1-24.

PYLYSHYN, Z. 1975. Do we need images and analogs? TINLAP-1, 174-177.

PYLYSHYN, Z. 1978. Imagery and artificial intelligence. In: Savage, C.W. (ed.), Perception and cognition: issues in the foundation of psychology. University of Minneapolis Press.

PYLYSHYN, Z.W. 1981. The imagery debate: analog media versus tacit knowledge. In: Block (1981).

PYLYSHYN, Z. 1984. Computation and cognition. MIT Press, Cambridge, Mass.

RUMELHART, D.E., NORMAN, D.A. 1980. Analogical processes in learning. CHIP 97, University of California, San Diego.

RUMELHART, D.E., NORMAN, D.A. 1983. Representation in memory. CHIP 116, University of California, San Diego.

RUSSELL, B. 1905. On denoting. Mind N.S. 14, 479-493.

SEARLE, J. 1980. Minds, brains, and programs. The Behavioral and Brain Sciences 3, 3, 417-457. Also in: Haugeland (1981).

SHEPARD, R.N., COOPER, L.A. 1982. Mental images and their transformations. MIT Press, Cambridge.

SIMON, H.A., NEWELL, A. 1956. Models: their use and limitations. In: L.D. White (ed.) The state of the social sciences, 66-83, University of Chicago Press.

SLOMAN, A. 1971. Interactions between philosophy and artificial intelligence: the role of intuition and non-logical reasoning in intelligence. IJCAI-2, 270-278. London. Also in: Artificial Intelligence 2, 209-225.

SLOMAN, A. 1975. Afterthoughts on analogical representations. TINLAP-1, 178-182.

SOWA, J.F. 1983. Conceptual structures. Information processing in mind and machine. Addison-Wesley.

STICH, S. 1984. Folk psychology and cognitive science: the case against belief. MIT Press, Cambridge, Mass.

WEYHRAUCH, R.W. 1980. Prolegomena to a theory of mechanized formal reasoning. Artificial Intelligence 13, 133-170.

WINOGRAD, T. 1975. Frames and the declarative-procedural controversy. In: Bobrow and Collins (1975).

WINOGRAD, T. 1983. Language as a cognitive process. Vol. 1 Syntax. Addison-Wesley.